

## Dynamic strategy selection in collaborative spatial tasks

Galati, Alexia; Panagiotou, Elisavet; Tenbrink, Thora; Avraamides, Marios N.

### Discourse Processes

DOI:

[10.1080/0163853X.2017.1293482](https://doi.org/10.1080/0163853X.2017.1293482)

Published: 01/08/2018

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

*Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):*

Galati, A., Panagiotou, E., Tenbrink, T., & Avraamides, M. N. (2018). Dynamic strategy selection in collaborative spatial tasks. *Discourse Processes*, 55(8), 643-665.  
<https://doi.org/10.1080/0163853X.2017.1293482>

#### Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Dynamic Strategy Selection in Collaborative Spatial Tasks

Alexia Galati<sup>1, 2</sup>

Elisavet Panagiotou<sup>1, 3</sup>

Thora Tenbrink<sup>4</sup>

Marios N. Avraamides<sup>1, 5</sup>

<sup>1</sup> Department of Psychology, University of Cyprus, Nicosia, Cyprus

<sup>2</sup> Cognitive and Information Sciences, University of California Merced, USA

<sup>3</sup> London School of Economics, London, UK

<sup>4</sup> School of Linguistics and English Language, Bangor University, Wales, UK

<sup>5</sup> Centre for Applied Neuroscience, University of Cyprus, Nicosia, Cyprus

Correspondence:

Alexia Galati  
Cognitive and Information Sciences  
University of California, Merced

agalati@ucmerced.edu

## Abstract

When speakers coordinate with one another, they have available a range of alternatives for conceptualizing and describing spatial relationships. To understand the features of successful communication in collaborative spatial tasks, it's important to identify factors that shape speakers' linguistic choices and evaluate them in relation to task success. In this paper, we examine how description strategies—in particular, references to global vs. local conceptualizations of spatial relationships—change over time, how the use of these strategies is related to both contextual cues and the partner's feedback, and finally how these factors affect communicative success in terms of efficiency and accuracy in the task. In the dialogue task we used, Directors described from memory a spatial layout with intrinsic properties to a Matcher who reconstructed it based on those descriptions. We found that global description strategies and feedback from the Matcher that contributed to grounding (such as recaps) predicted better task performance, whereas local description strategies and queries from the Matcher predicted poorer performance. Importantly, the strategy adopted by pairs early in the dialogue predicted their ultimate accuracy in reconstructing the layout. This work underscores that in order to unpack the complex factors that contribute to successful communication, it's important to consider not only the linguistic strategies that speakers use, but also how these unfold over time and are shaped by interactive processes, such as those reflected by the partner's feedback.

## Introduction

In everyday situations, people have to coordinate with one another in a variety of tasks that involve spatial thinking. Whether the task involves giving or following route directions, jointly assembling a piece of furniture, or searching the house for a misplaced object, interlocutors have available alternative ways for conceptualizing and, therefore, describing spatial relationships (e.g., Levinson, 2003). For instance, when providing directions, speakers may adopt the reference system of a navigator moving through the environment (e.g., “At the post office, turn right”, a so-called *route* perspective) or a reference frame that is allocentric and external to the environment (e.g., “At the post office, head east”, a so-called *survey* perspective) (Taylor & Tversky, 1996). In other tasks, in which conversational partners are physically co-present and act on their contingent environment, speakers may describe the location of objects relative to their conversational partner (“Please give me the bolt on your left”) or relative to themselves (“Please give me the bolt on my right”), among other options (Taylor et al., 1999).

Given this multiplicity of options for describing space, an important undertaking toward understanding features of successful communication is to identify some of the factors shaping speakers’ linguistic strategies, and to evaluate them in relation to task outcomes. Determining how to best describe a spatial relationship requires not only the coordination of cognitive processes within the mind of the speaker (i.e., perceiving and conceptualizing spatial relationships, planning spatial descriptions, and articulating them) but also interpersonal coordination across interlocutors (Brennan, Galati, & Kuhlen, 2010; Clark & Wilkes-Gibbs, 1986; Kraut, Lewis, & Swezey, 1982; Shockley, Richardson, & Dale, 2009). To understand the processes guiding the selection of spatial strategies, it is therefore important to study them in dialogue, the natural site of language use (e.g., Clark, 1996). Although the principles guiding the negotiation of perspectives in dialogue have been

explored to some degree (e.g., Carletta et al., 1997; Clark & Wilkes-Gibbs, 1986; Tenbrink et al., 2013), much less is known about the specific factors and processes associated with success on task outcomes (with some exceptions, e.g., Gergle, Kraut, & Fussell, 2004; Brennan, 2005). This could be because many naturalistic dialogue settings lack a basis for measuring success objectively.

However, spatial joint tasks often lend themselves well to assessing task success, since task goals are typically concrete (e.g., guiding the interaction partner to a specific destination or instructing them to place objects at specific locations) and thus afford a metric comparison of the actual end state of the task against the intended end state. Such tasks typically require the interacting partners' close coordination, since the availability of multiple strategies and potential conceptualizations requires them to establish "mutual knowledge of both conception and language" (Garrod & Anderson, 1987). Perceptual information, spatial viewpoint, prior spatial knowledge, or specific aspects of the task may differ for each partner, leading to a constant need to exchange information about the description scheme (and underlying spatial representation) being used in order to achieve coordination (e.g., Anderson et al., 1991, Garrod & Anderson, 1987).

In this paper, we focus on description strategies that involve conceptualizing spatial relationships in terms of a global reference frame (i.e., when spatial relationships within an array are integrated in an overarching representation) or in terms of a local reference frame (i.e., when the focus is on isolated spatial relationships within the array). We examine how speakers dynamically adapt their use of these global and local strategies not only in response to contextual factors serving as task constraints, but also in response to the partner's feedback. Critically, we also evaluate the contribution of these strategies to objective measures of task success.

### **Adapting spatial descriptions in response to contextual factors**

Speakers adapt what they say based on many different sources of information (Brennan, Galati, & Kuhlen, 2010). In terms of general, 'top-down' cues, speakers can consider their prior knowledge, beliefs, or expectations about their conversational partner, as well as other contextual cues that are perceptually available in the dialogue setting. In addition, they take into account 'bottom-up' cues that become available moment-by-moment as the conversation unfolds, including feedback from the addressee that reflects their engagement and understanding (Clark & Wilkes-Gibbs, 1986; Kraut, Lewis, & Swezey, 1982; Richardson, Dale, & Tomlinson, 2009).

Much of the research examining how top-down factors influence speakers' conceptualizations and descriptions of space has used monologic tasks in which speakers produce or interpret spatial descriptions in the absence of a contingently interacting partner (Mainwaring et al, 2003, Taylor et al., 1999). In a few dialogic studies, speakers have been shown to adapt their descriptions of space according to top-down information or attributions formed about their conversational partners. For instance, when speakers perceive their partner's ability to contribute to the task as more limited, they are more likely to adopt their partner's spatial viewpoint or elaborate their spatial descriptions. This is the case, for example, when the partner does not share their viewpoint (Schober, 1993, 1995) or has worse spatial abilities than they do (Schober, 2009). Such attributions about the partner influence the interpretation of spatial instructions as well. In a study in which listeners interpreted instructions that were ambiguous in certain visual contexts (e.g., "Give me the folder on the left" in a context where "left" could be interpreted as either the speaker's or participant's perspective), their beliefs about the speaker (whether they believed that the speaker was real vs. not, or whether the speaker knew their perspective vs. not) influenced their perspective choices (Duran, Dale, & Kreuz, 2011). Critically, beyond their distribution of perspective

choices, attributions about the partner also influenced the listeners' cognitive dynamics leading up to that choice, as reflected by their mouse trajectories to objects (e.g., deviations of the cursor towards the competitor folder, switches in the cursor's direction, and "acceleration components" involving the slowing down and then speeding up, taken to reflect hesitation). These micro-behavioral measures revealed that listeners co-activated spatial perspectives, as indicated for example by interference of the egocentric perspective during partner-centered responding or by the automatic activation of the partner's perspective in cases when it could have been ignored (e.g., in trials in which the object selection was the same from either perspective).

In addition to partner-specific factors, when producing or interpreting spatial descriptions, speakers also take into account environmental factors that can influence the interlocutors' assessment of their relative cognitive demands in the task and their subsequent description choices (Galati & Avraamides, 2013). Speakers can exploit features of objects that are perceptually available, such as their intrinsic axes, to override an otherwise prevalent egocentric description perspective (Tenbrink, Coventry, & Andonova, 2011). Speakers also use features of the environment, such as the presence of landmarks or prominent streets, incorporating them in their route descriptions with added details when addressing people unfamiliar with the environment (Hölscher, Tenbrink, & Wiener, 2011).

In the present work, we are interested in how both social and environmental cues influence how speakers adapt their use of global and local strategies when coordinating with their partners. When using a *global* spatial system, representations of spatial relationships are integrated within or subsumed by an overarching, global representation. In contrast, when using a *local* system, representations of local relationships and their connections are represented and recruited for spatial reasoning. Research on spatial memory—involving non-social, nonlinguistic experiments—has suggested that global reference frames are especially

useful to spatial reasoning, since tasks such as pointing, shortcutting, and estimating distances, rely on integrating local relationships into a higher-level global representation (McNamara, Sluzenski, & Rump, 2008). At the same time, local reference systems have also been shown to play an important role when making spatial judgments (Meilinger, Riecke, & Bühlhoff, 2014). Here, we examine the use of global and local systems in spatial descriptions, since to our knowledge their relative effectiveness has not been assessed in communicative settings.

### **Adapting spatial descriptions in response to the partner's feedback**

In addition to the top-down influence of contextual social and environmental cues, speakers' linguistic strategies are also influenced by bottom-up cues that become available moment-by-moment in the physical environment, such as those derived from the conversational partners' verbal and nonverbal behavior, including their progress on the task. As speakers present their utterances, they monitor their conversational partners for evidence of uptake and understanding (Brennan, 2005; Clark & Brennan, 1991). For example, speakers monitor their partner's eye-gaze to gauge what or whom their partners are attending to, and adapt accordingly their utterance planning or interpretation (e.g., Goodwin, 1979; Hanna & Brennan, 2007; Kendon, 1967). Thus, interlocutors continually seek and provide evidence about what they and their partners have understood, engaging in so-called *grounding* (Clark & Brennan, 1991; Brennan, 2005; Clark & Wilkes-Gibbs, 1986). In this way, the partner's feedback is essential to the dynamic development of the interaction.

The contribution of the partner's feedback has been studied increasingly in a variety of conversational settings, including storytelling (Bavelas, Coates, & Johnson, 2000; Pasupathi, Stallworth, & Murdoch, 1998; Kuhlen & Brennan, 2010) and referential communication (Horton & Gerrig, 2002), with some investigations in the spatial domain as well. In a spatial task in which dyads solved a modified version of the "maze game" (Garrod & Anderson,



1987) through a chat tool interface, Mills and Gregoromichelaki (2010) manipulated feedback by having clarification requests artificially introduced by the server at different points in the task as coming from the partner. Speakers interpreted clarifications requests differently, depending on whether they came earlier in dialogue or later. In earlier games clarification requests were taken to query the referential import of specific constituents in the speaker's previous turn, whereas in later games—by which point participants had become experienced and solved the game through highly elliptical exchanges—clarification requests were interpreted as questioning the purpose of the speaker's previous turn, indicating better intention recognition over time. Nevertheless, it's still not clear when such queries (and other types of feedback) are spontaneously deployed by naïve partners in joint spatial tasks, and how they would shape the speakers' strategy use and ultimate performance.

In this work, we systematically examine associations between different *types* of contributions from the partner (e.g., queries, proposals for reconceptualization, recaps) with linguistic strategies and task outcomes. Moreover, insofar as the partner's feedback is an ongoing source of evidence about their understanding throughout the interaction, we also seek to examine how the partner's feedback influences strategy use *over time*.

### **Adapting spatial descriptions over time**

Studies examining coordination in collaborative spatial tasks have typically focused on how distributions of spatial expressions change under different circumstances. Analyzing the total distributions of spatial expressions is useful in terms of capturing the speakers' aggregated preference in a particular description strategy, but it does not capture the incremental process by which strategies emerge and are negotiated in dialogue. This approach overlooks the possibility that such distributions may change over time (i.e., suggesting a change in description strategy) or may be shaped by the partner's feedback. Indeed, speakers have been shown to often mix perspective strategies (Tversky, Lee, & Mainwaring, 1999) and

to consider the partner's feedback and progress on the task to adapt those strategies (Schober, 2009).

Strategy choices may remain constant over time throughout an interaction, as in cases where partners converge on a particular conceptualization. For instance, as speakers repeatedly refer to the same potentially ambiguous objects, they reuse the same terms to signal that they are taking the same perspective to talk about the same entity, creating a "conceptual pact" (Brennan & Clark, 1996).

But strategy choices may also change during the course of the interaction, as when speakers have to update their attributions of their partner in response to the partner's feedback. For example, as speakers accrue knowledge about their partner's background, they adapt their descriptions accordingly. In a study by Isaacs and Clark (1987), speakers describing New York City landmarks were more likely to use more detailed descriptions (e.g., "that big building on the left") and less likely to refer to landmarks by their proper names (e.g., "the Chrysler building") when interacting with partners unfamiliar with New York City (novices) than NYC natives (experts). Importantly, whereas experts over time used consistently high proportions of proper names for landmarks to describe them to other experts, they decreased their use of proper names when describing them to novices. Similarly, in a study in which speakers interacted with a remote partner who they believed was either human or a computer that could interpret natural language, speakers over time adapted their descriptions to their partner's actual behavior (using more telegraphic turns or more complete sentences), despite their initial expectations (Brennan, 1991). These studies demonstrate that, over the course of the interaction, bottom-up cues from the partner can update or revise top-down expectations about the partner's ability to contribute to the task.

Speakers have been shown to adapt their spatial descriptions over time in collaborative spatial tasks as well. In a study by Schober (2009), partners were preselected to have matched

or mismatched spatial abilities, assessed by a mental rotation task, and to perform a spatial reconstruction task together. Speakers with high spatial ability were overall more likely to describe spatial relationships from their partner's viewpoint, whereas low ability speakers were more likely to use egocentric descriptions. Importantly, for mixed-ability pairs, high spatial ability speakers used more descriptions from the partner's viewpoint over time when interacting with low ability partners, whereas low spatial ability speakers used more egocentric descriptions over time when interacting with high ability partners. These findings further underscore that speakers form and update attributions about their partner's relative knowledge and ability based on unfolding cues provided by their partner, and that they jointly converge on a strategy—whether implicitly or explicitly—that is thought to be effective.

Such a strategy often requires the person with greater knowledge or ability to expend greater effort (e.g., adopting their partner's viewpoint vs. their own), in order to promote mutual understanding while minimizing collective effort (as posited by *the principle of least collaborative effort*, Clark & Wilkes-Gibbs, 1996). Whether the adaptation of strategies over time in response to the partner's feedback is in fact effective—in terms of improving coordination efficiency and task accuracy—is a separate empirical question, addressed next.

### **Assessing how spatial descriptions influence communicative success**

Obtaining success in task-related dialogic interaction involves a complex interplay of factors (Tenbrink et al., 2013). For instance, contextual factors such as the inter-visibility of interlocutors (and by extension their visual access to each other's actions and nonverbal feedback) can influence efficiency in the task. When one partner has visual evidence about what the other understands, pairs go through a shorter process of verbally checking that they mutually understood each other, compared to when such visual evidence is lacking (Brennan, 2005; Clark & Krych, 2004). Similarly, partners can coordinate more easily and achieve better task performance when they are co-present (versus not), especially in tasks that are

dynamic (e.g., with objects that are changing and are hard to describe), presumably because they are better able to apprehend the current state of the task (Gergle, Kraut, & Fussell, 2004). These findings suggest that in “more difficult” situations, pairs require a greater degree of *grounding* or exchanging of evidence about what they do or do not understand, giving rise to less efficient dialogues (e.g., Clark & Brennan, 1991; Clark, 1996).

Importantly, to reach the mutual belief that they have understood each other well enough for their current purposes, pairs must adopt a task-dependent criterion. This “grounding criterion” depends both on the affordances of the communicative situation (e.g., visibility between partners) and the goals of communication (Clark & Brennan, 1991). In fact, the goals of many spatial tasks typically prioritize accuracy over speed: it’s more important to guide your conversational partner to the correct destination or to assemble a shelf together correctly than to complete these tasks quickly. Thus, the spatial strategies used may have a differential impact on task outcomes capturing accuracy and efficiency, depending on the relative weighing of these goals in the task at hand. This is why, in this work, we examine how spatial strategies and the partner’s feedback relate to measures of both the efficiency of coordination and task accuracy.

### **Goals and approach**

Following the considerations laid out in the previous sections, we addressed four threads of inquiry concerning speakers’ description strategies in joint spatial tasks: (a) *Description strategies as a function of contextual cues*: how do speakers’ overarching description strategies (use of global vs. local reference systems) relate to top-down contextual cues that are available in advance, (b) *Description strategies as influenced by the partner*: how do these description strategies relate to bottom-up information, as reflected by the distribution of different types of partner’s feedback, (c) *Description strategies over time*: how do strategy choices change as the dialogue unfolds and in response to the partner’s feedback,

and (d) *Description strategies and task success*: which descriptions strategies are more successful in terms of improving the efficiency and accuracy on the task.

We addressed these questions by reanalyzing transcripts obtained from a study by Galati and Avraamides (2015), in which pairs of participants jointly reconstructed a spatial layout. In that work, the spatial expressions of the speaker describing the spatial layout (the Director) were classified as reflecting a particular “perspective”. This perspective represented an observer’s spatial point of view, in terms of the position of an axis relative to a relatum (Tenbrink, 2007), with the spatial axis, the relatum, and the origin being components of a spatial reference system. For example, the phrase “in front of me is the marble” involves having the Director as the relatum and was thus classified as Director-centered, whereas the phrase “the vase is to your left” has the partner (the Matcher) as the relatum and was classified as Matcher-centered. Galati and Avraamides (2015) found that speakers’ overarching preference for a particular spatial perspective depended on the available social and environmental cues, which were manipulated as described in more detail below. For example, speakers used partner-centered expressions more frequently when the partner’s viewpoint was aligned with the bilateral axis of symmetry of the configuration being described.

Although informative, these analyses in Galati and Avraamides (2015) do not capture the incremental process by which perspective preference emerges, or the dialogic negotiation elements involved. Neither do they address the broader description strategies in which spatial expressions participate. Notably, the majority of spatial expressions (48% of 1609) in Galati and Avraamides (2015) were from a “neutral” perspective capturing inter-object relations independent of a particular viewpoint (e.g., “it’s *close* to the bucket” or “they form a *triangle*”), whose overall distribution was not found to depend on the social and environmental factors manipulated in the study.

Here, we considered how such “neutral” spatial expressions (e.g., references to “lines”, “rows”, “columns”, “triangles” and other shapes) participate in broader conceptualizations of the layout, as part of a *global* or *local* reference system. As indicated above a global system is one that takes into account the layout’s structure, including most of its items (e.g., a system involving axes or a tic-tac-toe grid for conceptualizing, see Figure 1). A local system, on the other hand, is one that takes into account only isolated spatial relationships at any given time and makes reference to subsets of objects (e.g., conceptualizing the same figure in terms of isolated lines). Identifying these broader strategies enables us to tap into aspects of linguistic adaptation that the original study may have missed.

The methodological approach we used was that of Cognitive Discourse Analysis or CODA (Tenbrink, 2015), which involves the systematic parsing of dialogues into appropriate discourse units and the coding of theoretically motivated dimensions of content within those units, as well as associated linguistic and conceptual features. Dialogues are subjected to linearization, which permits examining over time the evolution of strategies and the interactive processes that support coordination (e.g., the partner’s feedback). Following the CODA approach, after using syntactic and prosodic cues to segment and linearize transcripts into appropriate discourse units for analysis, we coded each unit for the presence of a strategy of conceptualizing the layout as a global system or local system, and coded all discourse units in the partner’s (the Matcher’s) turns for their type of contribution (e.g., acceptance, query, recap, etc.). To address the effectiveness of global and local strategies, we examined how they related to measures capturing the efficiency of coordination (the number of conversational turns and discourse units, reflecting the length of the dialogue) and measures capturing accuracy on the task (derived from the final reconstructions of the configuration). By analyzing these non-linguistic measures from the tabletop reconstructions, along with

discourse measures derived from CODA, we aimed to triangulate the interactive processes governing successful coordination in spatial tasks.

## Method

### Description of original study yielding dialogues

The dialogues from Galati and Avraamides (2015) came from 24 pairs of Directors and Matchers (6 female-female, 6 male-male, and 12 were mixed-gender pairs, half with female Directors) who jointly reconstructed a spatial layout. After studying the layout of seven objects shown in Figure 1, the Director had their memory of the layout tested<sup>1</sup> and then described the layout from memory to the Matcher, who reconstructed it at a separate workstation based on those descriptions.

The study manipulated (1) the alignment of the layout's axis of symmetry (henceforth, its intrinsic structure) with either partner during the description phase, and (2) the Director's advance knowledge of their Matcher's viewpoint. In a third of the pairs, Directors studied the layout while aligned with its intrinsic structure (referred to as  $0^\circ$ ), and later described it to Matchers who were offset by  $135^\circ$  measured counterclockwise (*Aligned with Director* condition). In another third of the pairs, Directors studied the layout from  $225^\circ$  and later described it to Matchers who were at  $0^\circ$  (*Aligned with Matcher* condition). In the final third of the pairs, Directors studied the layout again from  $225^\circ$  and later it described to Matchers who were offset by  $135^\circ$ ; thus both partners were misaligned with the structure (*Aligned with Neither* condition). Half of the Directors in each condition studied the array while knowing their Matcher's viewpoint in advance (with the Matcher present in the room, seated at the position they would later occupy during the description phase), whereas the remaining half did not know the Matcher's subsequent viewpoint (the Matcher was absent from the room during the study phase).

During the description phase of the experiment, participants sat at separate tables, at the positions prescribed by their condition of alignment with the array's intrinsic structure. Pairs could interact freely but were separated by a barrier, such that they could see each other's faces but not each other's tabletops. Pairs were instructed to reconstruct the layout so that, given the Director's study viewpoint, the objects could be translated to the Matcher's table (i.e., not rotated by the Matcher's offset). The sessions were videotaped by two cameras, each with a view of one of the participants and their workstation.

[INSERT FIGURE 1 ABOUT HERE]

### **Transcripts**

The description phase for each pair was transcribed in detail, including annotations of fillers ("em" and "ee", Greek-Cypriot equivalents of "um" and "uh"), pauses, interruptions (both self-interruptions and interruptions by the partner) and restarts. The participants' non-verbal feedback, such as head nods and facial displays were also annotated in the transcripts, where relevant (see below).

### **Preparation of Transcripts for CODA analysis**

To linearize the transcripts, we first segmented the dialogues into *conversational turns*, and then identified within each turn its *basic discourse units* (Degand & Simon, 2009).

**Turns.** An uninterrupted stretch of speech by a Director or a Matcher was counted as a conversational turn. For our purposes, head nods that were unaccompanied by speech were also counted as turns only when they were viewed by the conversational partner over the barrier. Conversational turns are thought to reflect the pair's degree of *grounding*, or exchanging of evidence about what is and isn't understood (e.g., Clark & Brennan, 1991). As such, a decrease in the number of turns is thought to reflect facilitation in grounding due to successful coordination strategies that reduced one or both partners' cognitive cost of perspective-taking (e.g., Clark & Wilkes-Gibbs, 1986).



**Basic Discourse Units (BDUs).** Within each conversational turn, we identified BDUs based on both prosodic cues and the syntactic structure of the utterance. Following Degand and Simon (2009), prosodic and syntactic units were identified independently in separate passes. First, utterances in the transcripts were segmented into major intonation units by using perceptually detected prosodic features of the acoustic signal (i.e., by referring to the recordings of the dialogue), including its intonational contour, pauses, and the lengthening of the last syllable of the utterance (see examples in Table 1). Next, the same utterances in the transcripts were manually segmented into syntactic units, by identifying “dependency relations” between clauses (see Appendix A for more details about that segmentation process).

Once both intonation units and syntactic units were segmented in the transcripts, BDUs were identified based on the convergence of syntactic and prosodic boundaries. If only a prosodic or syntactic boundary was detected, the BDU continued until boundaries of both types coincided, as indicated by the numbering of BDUs in Table 1. That is, whether a syntactic unit constituted a BDU depended on the intonation contour overlaying that segment of speech. In addition to those BDUs determined by the convergence of prosodic and syntactic boundaries, utterances that were self-interrupted or interrupted by the conversational partner were also identified as standalone BDUs.

According to Degand and Simon (2009), by capturing syntactic and prosodic completeness, BDUs are thought to contain all the information necessary to support addressees’ inferences towards a coherent mental representation, contributing to the updating of that representation. By analyzing discourse at a level of granularity finer than turns, through BDUs, we can capture discourse content in informative ways that are not afforded by turns. For instance, the number of turns can serve as a proxy for the amount of back and forth between interlocutors without regard to how much is said within the speakers’ contributions.

BDUs, on the other hand, can capture the amount of informational content conveyed by speakers, thus distinguishing between single unit vs. multi-unit turns (e.g., a brief acknowledgement vs. a lengthy description) and permitting the coding of multiple content types that occur over time within a turn. The need for a fundamental structural distinction between turns and meaningful units within turns is widely recognized in the literature (e.g., Carletta et al., 1997; Schober, 2009).

[INSERT TABLE 1 ABOUT HERE]

### **Content Analysis**

Each BDU involved the following coding decisions: (a) whether it entailed a *global* or *local* strategy (in BDUs in both the Director's and Matcher's turns), and (b) if the BDU was part of the Matcher's turn, what type of feedback contribution it involved.

#### **Use of global vs. local systems.**

**Global systems.** When using a *global system*, speakers proposed a conceptualization that included most or all of the items of the layout in order to describe their relative position. For instance, one Director proposed conceptualizing the layout as an "X-O grid" (the Greek conventional terms for the game of *tic-tac-toe*), which involved imagining a 3-by-3 grid superimposed on the table. This global strategy enabled the pair to number the nine boxes and coordinate the placement of objects in seven of those boxes. Other examples of global systems involved conceptualizing the layout as clock (with its hours as directional reference points), or as forming the shape of a house, a cross, or a system of axes and quadrants.

A related sub-category, which we included in the count of global references, involved reference to the system's constituent elements (*global constituents*). For instance, reference to an individual numbered box or an individual row of boxes of the aforementioned "X-O grid" system was coded as a reference to constituents of the global system.

**Local systems.** When using a *local system*, speakers recruited only a small subset of the items (up to three) to describe the relative positions of objects. Examples of using a local system included references to the bucket as the center of the layout, descriptions of a small number of objects forming a geometric shape (e.g., the flashlight, yoyo, and battery forming a right-angle triangle), or descriptions of lines (e.g., D4: “on the same line as the bottle and the marble, next there will be the\*... the candle ”) or of objects forming lines (rows, columns, or diagonals) without the mention of an overarching global system (e.g., the “X-O grid”).

Notably, a reference to lines, columns, rows, or diagonals (comprising two or three objects) could be coded either as *local* or as a constituent of a global system (*global constituent*), depending on the preceding and following dialogue. Based on the surrounding discourse context, the coders assessed whether the description of such a line referred to part of a global system proposed by either partner and made their judgments accordingly. For example, D1’s reference to “three horizontal lines” in Table 1 was coded as *global constituents* not only because it captured multiple objects at once (6 in total) but also because it was followed by additional references to horizontal and vertical lines (as perceived from D1’s perspective of 225°) that formed a global system; subsequent references to a single horizontal line, were also coded as a *global constituents*.

**Partner’s feedback.** Each BDU within of a Matcher’s turn was coded for the type of contribution it made. We expanded the categories described by Horton and Gerrig (2002), which resulted in the following coding scheme:

1. *Acceptance*: the Matcher indicated successful uptake of the Director’s description, typically through an affirmative response such as “yes”, “got it”, “I understood”, or “OK”. It also included cases in which the Matcher interrupted the Director’s description to indicate acceptance of their description.

2. *Query: Clarification request:* the Matcher requested clarification of some portion of the Director's previous description (e.g., asking "how can it [the marble] be perpendicular to the yoyo?" after the Director had said "and take your marble, which must be perpendicular to the yoyo"), or otherwise by posing a yes/no question, e.g., "so to my left is the flashlight and then the battery?").
3. *Query: Expansion request:* the Matcher either implicitly requested an expansion of the previous description (often by using fillers such as "ee" or saying "yes?" with rising intonation) or explicitly did so (e.g., "hold on, though, how much distance?").
4. *Expansion: new proposal:* the Matcher proposed a novel expansion that was not part of the Director's earlier description (e.g., saying "does this mean that the bucket and the candle will be side by side?" after the Director had said "move vertically, and place the candle to the right of the yoyo", or saying "in other words, it formed something like a zigzag" when the Director's previous descriptions did not involve reference to a zigzag system).
5. *Response to question:* the Matcher responded to the Director's question (e.g., saying "directly in front of me, I have to have the marble", after the Director asked "let's do a verification, OK? Directly in front of me [what do I have], are you listening?").
6. *Description-recap:* the Matcher described the layout. Any description of spatial relationships that was not a response to the Director's question in the previous turn, was considered to belong to this category. Typically, these descriptions took place after the Matcher had placed all objects on their table, and described either part or the entirety of the layout. For example, in pair 1, M1 initiated a recap spanning 14 turns, after D1's description of the array; all BDUs by M1 in the following excerpt were coded as description-recap:

M1: we have the flashlight and the<e>

D1: yoyo /

M1: the yoyo

D1: yeah

M1: then behind that / in between the flashlight and the yoyo we have the  
bucket

7. *Metacomment on task*: These contributions pertained to progress of the task and other aspects of the interaction, as opposed to referring to the spatial configuration itself. Metacomments included contributions in which the Matcher expressed confusion, understanding, or apology for error, or negotiated with the Director the reference frame or type of system used; contributions pertaining to the management of the task (e.g., announcing or agreeing on the conclusion of the task, or indicating the need to restart a segment of the description) or to task rules (e.g., not looking over the barrier); comments on their perceived success on the task (e.g., “we got it!”) or the effectiveness of themselves or their partner (e.g., “you’re a God!”).
8. *Uncodable*: Contributions in which the speech was unintelligible, or else interrupted such that the remaining speech fragment did not permit a coding judgment.

Table 2 includes samples of coding from two pairs, illustrating how strategy use (global and local references) and the Matcher’s contributions were coded in dialogues that have been linearized by BDUs.

[INSERT TABLE 2 ABOUT HERE]

**Reliability.** The second author coded all pairs, while another coder redundantly and independently coded BDUs in linearized transcripts from six pairs (approximately 25% of the corpus) for the use of strategies involving a global system, a local system, and the Matcher’s feedback (in the Matchers’ BDUs). Levels of agreement were very good: Krippendorff’s alpha (computed with the macro reported in Hayes & Krippendorff, 2007) for identifying a

global strategy was .93, for identifying a local strategy was .80, and for classifying the Matcher's feedback (as acceptance, query, new proposal, response to question, recap, metacomment, and uncodable) it was .89.

### Measures of task success

**Communicative efficiency.** We assessed communicative efficiency in terms of the numbers of turns and BDUs produced by pairs, which we took to reflect the length of the dialogue.

**Task accuracy.** In order to determine task accuracy, we considered how accurately Matchers reconstructed the configuration. For every reconstructed array, we had taken a bird's-eye view digital photograph and by superimposing a grid we extracted the coordinates of the layout's seven objects, comparing them to those of the original layout through bidimensional regression analyses (Friedman & Kohler, 2003). Reconstruction accuracy was assessed through two measures: (i) the bidimensional regression coefficient ( $BDr$ ), and (ii) the rotation parameter ( $\theta$ ). In bidimensional regression analyses, a Euclidean transformation is applied to the set of 7 dependent A-B points (corresponding to the Matcher's placement of the 7 objects on the table), such that they are optimally rotated, scaled, and translated to match the 7 fixed independent X-Y points (corresponding to the veridical coordinates of the objects that the Director had studied, shown in Figure 1). The adjusted points are then correlated with the correct response, resulting in a correlation coefficient ( $BDr$ ), which estimates the goodness-of-fit between the reconstructed and the actual coordinates of the layout, thus capturing unsystematic error in reconstructions when systematic biases are accounted for. The rotation parameter ( $\theta$ ) indicates the degree to which tabletop reconstructions were rotated relative to the studied layout, thus capturing a potential systematic bias in the reconstructions.

## Results and Discussion

### I. Description strategies

**Use of global and local strategies as a function of social and environmental factors.** Overall, pairs referred to a global system or its components in 9.17% of all BDUs ( $SD= 8.45\%$ ) and to a local system in 6.86% of all BDUs ( $SD= 5.99\%$ ). The use of global and local systems tended to be complementary (i.e., as pairs used more global systems references they tended to use fewer local system references, and vice-versa), although this correlation was not significant, Pearson's  $r = -.37, p = .08$ .

Our examination of the relationship between strategy use and the experimentally manipulated social and environmental factors did not reveal any main effects or interactions in an ANOVA with strategy (global, local) as a within-pairs factor, and structure alignment and availability of the Matcher's viewpoint as between-pairs factors. Nevertheless, the patterns illustrated in Figure 2 suggest some differences in the distribution of global and local strategies across these contextual factors.

For instance, global strategies were more prominent than local strategies when pairs knew in advance that neither of them was aligned with the layout's intrinsic structure. As seen in the rightmost black bars for global and local strategies in Figure 2, in the Neither-aligned condition, pairs were more likely to use a global system than a local system when they knew in advance of their relative viewpoints,  $F(1, 18) = 4.90, p = .04, \eta^2 = .21$ . This could be because advance knowledge of their oblique viewpoints helped partners leverage the properties of the structure in their descriptions.

Global strategies were also more prominent when Directors were aligned with the intrinsic structure and didn't know at study where the Matcher would be. As seen in the leftmost white bars for global and local strategies in Figure 2, pairs tended to use more global than local references,  $F(1, 18) = 4.05, p = .06, \eta^2 = .18$ , when the Director was aligned with the

intrinsic structure but the Matcher's viewpoint was unavailable at study. This could be because, in the absence of information about the partner, the layout's global properties at study were highlighted from the Directors' vantage point.

Finally, local strategies were more prominent when the Matcher was aligned with the intrinsic structure and this was known in advance. When considering only the distribution of local strategies in the Matcher-aligned condition, pairs were significantly more likely to use local systems when the Matcher's viewpoint was available at study than when it was unavailable,  $F(1, 18) = 5.18, p = .04, \eta^2 = .22$ . Knowing in advance that the Matcher was aligned with the intrinsic structure may have motivated Directors to describe the configuration in a piecemeal fashion through references to local spatial relationships.

[INSERT FIGURE 2 ABOUT HERE]

## II. Strategy use and the Matcher's contributions

**Strategy use by the Matcher.** Overall, Matchers contributed on average 45% of the BDUs in the dialogue (range: 35-54%). In terms of strategy use, they contributed on average only 34% of the BDUs that contained references to a global system or its components ( $SD = 30\%$ ) and 31% of the BDUs that contained references to a local system ( $SD = 21\%$ ). That Matchers made fewer contributions than Directors was expected given the informational asymmetry of the task, whereby Directors had privileged information about the spatial layout having studied it previously.

Matchers produced fewer references to global and local systems in the Director-aligned condition compared to the other conditions, although these differences were not significant. Out of all BDUs with references to global systems in a given dialogue, Matchers produced on average 19% of them in the Director-aligned condition ( $SD = 10\%$ ) compared to 42% in each of the Matcher-aligned and Neither-aligned conditions ( $SD = 44\%$  and  $SD = 24\%$ , respectively). Similarly, out of all BDUs with references to local systems in a given dialogue,



Matchers produced on average of 21% of them in the Director-aligned condition ( $SD= 16\%$ ) compared to 35% in each of the Matcher-aligned and Neither-aligned conditions ( $SD= 31\%$  and  $SD= 10\%$ , respectively).

These quantitative patterns were in line with our observation that Matchers were more likely to take initiative to propose a spatial strategy, particularly a global one, when the Director was not aligned with the intrinsic structure. This point is highlighted in Appendix B, where the entire distribution of *unique* global strategies (i.e., global systems not previously introduced by the other partner), contributed by each speaker, can be seen across the 24 pairs. Of the 8 pairs in which Matchers introduced novel, unique global strategies, 4 pairs were in the Neither-aligned condition, 3 in the Matcher-aligned condition, and only 1 in the Director-aligned condition.

For illustration of such a case, consider the excerpt of dialogue in Table 2, in which a Matcher in the Neither-aligned condition proposed a global system. In this example, M1 proposes conceptualizing the objects as a zigzag shape—a global system that differs from the one previously proposed by D1 (which involved horizontal and vertical lines, as we’ve seen from the beginning of the dialogue in Table 1). D1 accepts M1’s conceptualization of the zigzag with some qualification. Echoing the quantitative trends above, this example instantiates a case where the Matcher likely appraised that they were in a good position (literally and figuratively) relative to their partner to apprehend the emerging spatial relationships in the reconstructed configuration and to contribute to the task by introducing a novel global strategy for conceptualizing the layout.

Although the distribution of the Matcher’s novel contributions (i.e., “new proposals” in our coding scheme) did not change significantly across conditions of structure alignment, the distribution of their queries (clarification and expansion requests combined) did change,  $F(2, 18)= 3.58, p = .049, \eta^2 = .29$ . This distribution was parallel to that of spatial strategies

described above, with Matchers posing fewer queries when Directors were aligned with the intrinsic structure (only 7.68% of the total BDUs) compared to the other two conditions (13% in the Matcher-aligned and 11.59% in the Neither-aligned conditions). The difference between the Director-aligned and Matcher-aligned condition was marginally significant, 95% CI [-.11, .011], Bonferroni-adjusted  $p = .056$ .

Thus, when Directors were at 0°, Matchers were less likely to refer to spatial strategies (global or local) and less likely to ask questions than in the other alignment conditions. These patterns could be because Directors, when aligned with the structure, provided clearer descriptions that required fewer interjections from the Matcher, or because Matchers (from their oblique viewpoint relative to their partner's aligned viewpoint) did not consider themselves to be well positioned to recruit spatial strategies or question their partner's descriptions.

**Strategy use and the Matcher's feedback.** When examining the relationship between the Matcher's feedback and the use of a global or local system, the following notable patterns emerged. First, as the Matcher's use of metacomments increased, the proportion of BDUs with global component references (contributed by either the Director or Matcher in a given dialogue) also increased, Pearson's  $r = .40$ ,  $p = .05$ , as the use of a global system may have required more management of the task.

Second, as the proportions of BDUs containing recaps by the Matcher increased, the proportion of BDUs in a given dialogue with local system references contributed by the Matcher also increased, Pearson's  $r = .46$ ,  $p = .02$ . The Matchers' use of local systems (i.e., the proportion of BDUs with local references contributed by the Matcher vs. the Director per dialogue) was also highly correlated with their posing of clarification requests (the proportion of BDUs classified as clarification requests per dialogue), Pearson's  $r = .80$ ,  $p < .001$ . Both of these patterns could be either because these types of contributions (recaps and clarification

requests) contained references to local relationships or because they co-occurred with the Matcher's increased use of a local system elsewhere in the dialogue.

### III. Changes in discourse over time

**Strategy use over time.** In order to examine how the pairs' strategies evolved over time, for each pair we selected the first 1/3, second 1/3, and final 1/3 of their BDUs, and computed the proportions of strategies occurring within each segment.

An ANOVA with dialogue segment (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>) and strategy use (global, local) as within-pair factors, and with structure alignment and the availability of the Matcher's viewpoint as between-pair factors, revealed a significant three-way interaction between dialogue segment, structure alignment, and the availability of the Matcher's viewpoint,  $F(4, 36) = 2.70, p = .046, \eta^2 = .14$ . This interaction is contextualized by the finding that, in the difficult condition in which neither partner was aligned with the layout's structure and the Matcher's viewpoint was unavailable in advance, pairs referred to global and local systems more frequently in the final third of the dialogue than in the earlier segments (3<sup>rd</sup> vs. 1<sup>st</sup> segment 95% CI [.01, .13],  $p = .03$ ; 3<sup>rd</sup> vs. 2<sup>nd</sup> segment 95% CI [.01, .13],  $p = .03$ ). That is, in this difficult condition conceptualizing the layout in terms of these strategies was more likely to emerge later in the dialogue. The example of pair 1 in Table 2 is one such data point, as references to a strategy (a global system) were introduced by M1 in the final segment of the dialogue (with BDU no. 253 corresponding to the 80<sup>th</sup> percentile of that dialogue's BDUs).

As illustrated in Figure 3, references to local and global strategies patterned differently over time. For local systems there was an increase in their use over time, which was parallel when the Matcher's viewpoint was available and unavailable, although this increase was not statistically significant when assessed through a linear contrast,  $F(1, 18) = 2.23, p = .15, \eta^2 = .11$ . By contrast, for global systems, there was a drop in their use in the final segment of the dialogue, but only when the Matcher's viewpoint was available in advance (linear contrast,  $F$

(1, 9) = 4.76,  $p = .057$ ,  $\eta^2 = .35$ ). When the Matcher's viewpoint had been unavailable at study, pairs continued referring to global systems with relatively high frequency in the final segment of the dialogue, with no significant differences across segments (linear contrast,  $p = .55$ ).

Earlier, when we described the complimentary relationship between the two description strategies, we noted that the negative correlation between the frequency of global and local references did not reach significance ( $p = .08$ ). Interestingly, when considering the relationship of these strategies over time, this correlation *was* significant in the beginning of the dialogue (the first 1/3 of BDUs, Pearson's  $r = -.52$ ,  $p = .01$ ), marginal in its middle segment (Pearson's  $r = -.36$ ,  $p = .09$ ), and non-significant in the final segment ( $p = .36$ ). That is, the complementarity of the two strategies (or the preference of one strategy over the other) attenuated as the task progressed.

[INSERT FIGURE 3 ABOUT HERE]

**Matcher's feedback over time.** One notable pattern observed in Figure 4, which illustrates how the Matcher's contributions across the three segments of the dialogue, is the sharp increase in recaps in the final segment of the dialogue. This increase makes sense in the context of a grounding strategy that involves having the Matcher re-describe the layout after reconstructing it, as many pairs did. The 3<sup>rd</sup> segment of the dialogue had significantly more BDUs with recaps compared to the 2<sup>nd</sup>, 95% CI [.032, .14], Bonferroni-adjusted  $p = .002$ , which in turn had more recaps than the 1<sup>st</sup> segment, 95% CI [.001, .034], Bonferroni-adjusted  $p = .03$ . A similar pattern was observed for metacomments, which increased in the final segment relative to the previous segments (3<sup>rd</sup> vs. 2<sup>nd</sup>: 95% CI [.014, .055], Bonferroni-adjusted  $p = .001$ , 3<sup>rd</sup> vs. 1<sup>st</sup>: (3<sup>rd</sup> vs. 2<sup>nd</sup>: 95% CI [ $6.90 \times 10^{-5}$ , .051], Bonferroni-adjusted  $p = .049$ ), indicating that there was additional need for task management as the task approached its conclusion.

By contrast, queries and new proposals from the Matcher decreased over time. Matchers made significantly fewer queries (clarification and expansion requests) across the three segments (linear contrast  $F(1, 18) = 31.52, p < .001, \eta^2 = .64$ ), and decreased their new proposals from the 2<sup>nd</sup> to the 3<sup>rd</sup> segment ( $F(1, 18) = 5.22, p = .04, \eta^2 = .23$ ). Collectively, the reduction over time of clarification questions, expansion requests, and new proposals makes sense because, as Matchers reconstructed more of the layout and pairs presumably converged on a description strategy, this type of feedback became less necessary.

[INSERT FIGURE 4 ABOUT HERE]

#### IV. Task success

**Task success and strategy use.** The use of global systems was, overall, weakly associated with increased accuracy on the task: as pairs used greater proportions of BDUs with global references, Matchers produced reconstructions that were less distorted (higher Fisher-transformed  $BDr$ ), Pearson's  $r = .34, p = .09$ , and less likely to be rotated (smaller the greater angle of rotation,  $\theta$ ), Pearson's  $r = -.36, p = .09$ . Importantly, the correlation between global references and  $BDr$  was marginally significant in the first segment of the dialogue (Pearson's  $r = .38, p = .08$ ), significant in the second segment (Pearson's  $r = .53, p = .01$ ), and not significant in the final segment ( $p = .76$ ), suggesting that global references were more beneficial earlier in the interaction than towards its end.

Conversely, as pairs used increasing proportions of BDUs with local references, Matchers tended to produce reconstructions that were more distorted (in terms of lower  $BDr$ ), Pearson's  $r = -.38, p = .08$ . Specifically, reconstructions were more distorted as local references came increasingly from the Matcher (i.e., the proportion of BDUs with local references in a given dialogue contributed by the Matcher vs. the Director), Pearson's  $r = -.58, p < .01$ . Again, this relationship was stronger earlier in the dialogue: the negative correlation between  $BDr$  and the proportion of BDUs with local system references was

significant in the first segment on the dialogue, Pearson's  $r = -.49$ ,  $p = .02$ , but non-significant in the subsequent two segments ( $p = .24$  and  $p = .21$ , respectively).

**Task success and the Matcher's feedback.** When considering how communicative efficiency related to the content of the Matchers' contributions, the main pattern that emerged was that with increasing proportions of BDUs with metacomments indicating the Matcher's confusion, the length of the dialogue increased (for turns: Pearson's  $r = .71$ ,  $p < .01$ , and for total BDUs, Pearson's  $r = .75$ ,  $p < .01$ ). This makes sense insofar as the Matcher's expressed confusion could have prompted some back and forth with the Director in an attempt to resolve it.

In terms of task accuracy, as Matchers made more recaps, their reconstructions were more accurate, involving a smaller angle of rotation, Pearson's  $r = -.46$ ,  $p = .03$ , and higher *BDr*, Pearson's  $r = -.37$ ,  $p = .08$ . Similarly, as Matchers responded to more of the Directors' questions, reconstructions tended to be more accurate, with higher *BDr*, Pearson's  $r = .39$ ,  $p = .07$ . Higher *BDr* was also associated with increased acceptances by the Matcher (Pearson's  $r = .38$ ,  $p = .08$ ), perhaps because acceptances reflected clearer descriptions from the Director. Collectively, layout reconstructions improved as the Matcher made contributions that facilitated the process of *grounding*, by ratifying agreement on object relationships through recaps, responses questions, and the acceptance of descriptions.

Interestingly, greater proportions of BDUs with queries from the Matcher (combined clarification and expansion requests) were associated with reconstructions that were more distorted (lower *BDr*), Pearson's  $r = -.61$ ,  $p < .01$ . Beyond the possibility of a direct pernicious effect of queries on performance, it's possible that a characteristic of the Matcher or the Director (e.g., their spatial ability, which could influence how well they interpreted or planned spatial descriptions) accounts for both the frequency of queries and the reconstruction performance. New proposals by the Matcher were also associated with greater angles of

rotation in the reconstruction (Pearson's  $r = .41$ ,  $p = .05$ ), suggesting that when the Matcher took initiative by reconceptualizing spatial relationships this was associated with worse performance.

[INSERT TABLE 3 ABOUT HERE]

### **General Discussion**

In this work we examined how spatial strategies (references to global vs. local systems) are recruited as a function of contextual cues (the alignment of a spatial layout with the interlocutors' relative positions, and their advance knowledge of that) how these strategies relate to the partner's contributions, how they develop over time, and how they predict success on the task. Table 3 provides a summary of our main findings across these four threads of inquiry.

Before addressing these points, it is worth underscoring the finding that global and local description strategies in this collaborative task had a complementary relationship: pairs who used more global references used fewer local references, and vice-versa. The complementary use of spatial strategies was more evident in the earlier segments of the dialogue, with their association becoming attenuated when the task approached its end. This may suggest that, initially, conversational partners converge on a particular system for conceptualizing the configuration at the expense of another (e.g., Brennan & Clark, 1996), but as they make progress on the task this preference attenuates and they may be increasingly inclined to use alternative conceptualizations of the configuration. This is consistent with Garrod and Anderson's (1987) observation that, even after explicit negotiation of a spatial scheme, conversational partners may not adhere to that scheme for the entire conversation.

This finding also resonates with other findings that people use diverse cognitive strategies to simplify complex spatial problems. One such domain is the "traveling salesperson problem" (TSP), in which the goal is to find the shortest way of connecting a

number of locations to each other before returning to the starting position. When solving this problem, people typically use a number of heuristics to reduce the problem's complexity, which they do with good and efficient results. Relevant to our present findings, problem solvers of the TSP have often been found to start out focusing on a coarser strategy (e.g., identifying object clusters or coarse trajectories) and subsequently refine their trajectory in detail to include individual targets (Graham, Joshi, & Pizlo, 2000; Tenbrink & Wiener, 2009). Similarly in our study, some pairs benefitted from establishing global strategies early on, but over time mixed those strategies with local ones to describe the more fine-grained spatial relationships of the layout. Others were more focused on local relationships at the start and then gradually opened up their cognitive scope towards the whole arrangement, similar to 'fine-to-coarse' heuristics in route planning (Wiener & Mallot, 2003). The interplay of global and local layers has also been recognized as relevant for other cognitive domains, such as perception (Förster & Higgins, 2005) or mathematical problem solving (Garofalo & Lester, 1985). Thus, global and local strategies don't necessarily involve mutually exclusive spatial schemas, but rather they can be recruited flexibly throughout a collaborative task.

In terms of our first research question, as summarized in Table 3, the use of global and local strategies was influenced to some degree both by the a priori availability of the partner's spatial perspective and by environmental cues (the structure's alignment). This is in line with earlier evidence that other kinds of linguistic choices of the same speakers (their use of egocentric or partner-centered expressions) also depended on both factors (Galati & Avraamides, 2015, 2013). The present work extends those findings by demonstrating that, when Directors were misaligned with the layout's structure, pairs were more likely to use a local strategy when knowing in advance that the Matcher would be aligned with the structure and were more likely to use a global strategy when knowing in advance that the Matcher would also be misaligned with the structure. This latter finding, in particular, suggests that



pairs in the difficult Neither-aligned condition (which required more turns to coordinate, cf. Galati & Avraamides, 2013) were more likely to leverage information about the structure's symmetrical properties when knowing the Matcher's viewpoint in advance.

In terms of the second research question, concerning the partner's contributions, we found that Matchers spoke less than Directors, contributing less than half of the total BDUs and about a third of the references to global and local systems (which occurred in about 15% of the total BDUs). This level of engagement makes sense in light of the informational disparity between partners in the task. Having studied the layout as a whole, Directors were better poised than Matchers to propose a system (whether global or local) to conceptualize the layout, and dominated the conversational floor (see also Tenbrink, Andonova, & Coventry, 2008). We found that, at least in a task with such informational asymmetry, certain types of contributions from the partner, such queries or new proposals, were associated with detriments in task performance; this relationship warrants further empirical exploration.

Critically, our findings underscore that partner's feedback shapes task success, consistently with a view of language use that regards addressees as co-creators and co-narrators in dialogue (e.g., Clark, 1996; Bavelas, Coates, & Johnson, 2000). As pairs engaged in more "grounding", by ratifying what was mutually understood through increased recaps, acceptances of the Director's proposals, and responses to the Director's questions, tabletop reconstructions became more accurate (see point 4b in Table 3). In contrast, increased queries were associated with less accurate reconstructions. Reconstructions were also more distorted as the partner took initiative, in the form of increasing new proposals or increasing proportions of local system references (generated by the Matcher vs. the Director).

Task success was predicted by the spatial strategies of interest as well: the use of global strategies was associated with increased success on the task as reflected by the accuracy of the tabletop reconstructions. As pairs used more global references, they tended to

produce reconstructions that were less distorted and less likely to be rotated. Again, these findings extend earlier work based on the same experiment that found no correlation task success and the Directors instructions in terms of their spatial perspective (i.e., speaker-centered, partner-centered, or structure-centered, Galati & Avraamides, 2013). The general benefit of global strategies, demonstrated here, resonates with research on spatial memory that suggests that representing spatial locations in a single global reference frame underlies the ability to do well in an array of spatial reasoning tasks (McNamara, Sluzenski, & Rump, 2008). Conversely, as pairs in our task used more local references, they produced reconstructions that were more distorted.

Importantly, our findings suggest that these associations held mainly for the earlier segments of the dialogue: increased global references were more beneficial to task accuracy and increased local references were more harmful to task accuracy when occurring earlier in the dialogue than towards its end. This point warrants emphasis: patterns in coordination can often be obfuscated when the distribution of linguistic choices is considered for the entire dialogue (treating it as a fixed corpus), but they may be unveiled when taking into account their development over time. Indeed, all correlations between strategy use and the measures assessing the Matchers' reconstructions were marginally significant for the dialogue as a whole, but became significant when focusing on earlier dialogue segments. In addition to this finding, other shifts in discourse over time can be seen in Table 3 (research question 3).

For instance, recaps and metacomments increased over time, whereas queries and new proposals for conceptualizing the layout decreased over time. These changes in feedback make sense in the context of the "grounding criterion" (Brennan & Clark, 1991) of this task (which emphasized accuracy in the reconstructions) and the task's affordances (e.g., the fact that partners could not see each other's respective work areas). With respect to recaps, Matchers often re-described the layout after reconstructing it, as a way of double-checking

object placements with the Director. And with respect to queries and new proposals, these likely became less necessary as more spatial relationships among objects were agreed upon in the process of reconstructing the layout. The declining frequency of queries over time is compatible with findings that queries are interpreted differently late in dialogue, once interlocutors are sufficiently coordinated, compared to early in dialogue (Mills & Gregoromichalaki, 2010). The level of analysis we employed here, which considers the content of the partner's feedback, extends prior work that has focused on the speakers spatial perspective choices over time, but without examining the type of contribution (or "illocutionary force" of the utterance) within which spatial expressions were embedded (e.g., Schober, 2009).

Our undertaking to quantify changes in discourse over time resonates broadly with that of a dynamical systems approach to cognition, which seeks to characterize through common principles the emerging behavior of complex systems (whether biological, cognitive, or social), with an emphasis on the timescales at play. This approach has been extended to domains that include high-level coordination in dialogue (e.g., Duran & Dale, 2014; Fusaroli, Raczaszek-Leonardi, & Tylén, 2013). For instance, a dynamical model of spatial perspective-taking, in which attributions about the conversational partner are represented as weighted information evolving over time, has been shown to account well for the motion dynamics of responses to spatial instructions (namely, the participants' mouse-trajectories within a given trial) as well as for the stabilization of perspective choices over time (Duran & Dale, 2014). Although we are operating at a coarser temporal grain here, based on informational units, the CODA discourse analysis approach we employ (Tenbrink, 2015) permits capturing, through the temporal linearizing of dialogues, systematic patterns in the linguistic choices of all conversational participants, and has the potential to reveal coordination patterns that unfold over time.

In sum, our findings provide some intriguing suggestions regarding how people coordinate in collaborative tasks. In this task, in which partners had to reconstruct a spatial configuration with intrinsic properties, queries from the partner were associated with the use of local systems, which was in turn associated with poorer task performance. In contrast, the use of global systems and feedback from the partner contributing to grounding (e.g., recaps) were both predictive of better task performance. Importantly, the kind of spatial strategy established early in dialogue set the stage for task success, with the early use of global systems predicting better accuracy and the early use of local systems predicting poorer accuracy. Future work that considers the unfolding of spatial strategies over time and their contingencies to the partner's feedback and their locally co-occurring actions, can afford a more nuanced understanding of what determines successful coordination in collaborative settings.

### **Authors' note**

This material is based upon work supported by the European Research Council under grant 206912-OSSMA to M.A. We are grateful to Mary Vasiliou for assistance with coding. We also thank David Rapp, Riccardo Fusaroli, and two anonymous reviewers for suggestions that significantly improved this paper.

## Appendix A

### Segmenting utterances into syntactic units

The utterances in the transcripts were manually segmented into syntactic units by identifying “dependency relations” between clauses, following Degand and Simon’s (2009) use of dependency syntax theory (see Heringer, 1993, for a review). In this theory, syntactic units consist of a “nucleus” (typically a verb) accompanied by its clauses, which may include complements governed by the verb, or adjuncts. Thus, a given syntactic unit contains clauses that are embedded within dependency relations. The constituency of clauses was determined by a process of pseudo-clefting, whereby each clause was shifted into focus at the beginning of the sentence: insofar as clauses could be clefted they were thought to be connected by dependency relations and were considered part of the same syntactic unit, as we illustrate below. According to our segmentation criteria, adjuncts headed by conjunctions such as “but”, “and”, “yet”, interjections such as “so”, as well as discourses markers such as “I mean”, “let’s say”, were included in the syntactic units that preceded or followed them (cf. Degand & Simon, 2009, who excluded these elements from the syntactic unit).

In example 1, bracketed elements were judged to be bound by dependency relations and to be contained within a single larger syntactic unit, marked by the larger square brackets. (Speech is transliterated from Cypriot Greek and translated in English underneath, with pauses indicated by slashes.)

Example 1:

D12: [ lipon /

[metaxi / siklas / dje tis plevras pu stekese esi

alla tis akrias tis kanonikis ] /

[katse to fanari / ] ]

[ so /

[between / the bucket / and the side that you are standing at

but the regular side ] /

[place the flashlight / ] ]

This larger syntactic unit is governed by the verbal clause “place the flashlight”, which has one main clause<sup>1</sup> dependent to it: “between the bucket and the side you are standing at, but the

---

<sup>1</sup> The clause “between the bucket and the side you are standing at, but the regular side” can be seen as comprising constituent sub-clauses: one headed by the conjunction “but” (“but the regular side”), and the two noun phrases bound in a dependency relation by the conjunction

regular side<sup>2</sup>). When using pseudo-clefting as the test for constituency, we confirmed that in Cypriot Greek, as in English, the two clauses could be transposed to bring the placement of the flashlight into focus (yielding the sentence “place the flashlight between the bucket and the side that you are standing at, but the regular side”). As we noted above, according to our segmentation criteria, the initial interjection “so” was grouped with the syntactic unit that followed it.

Although syntactic dependency clauses were typically governed by a verb and its governing complements, they could also be averbal, as in example 2 below, which was M12’s initial response, following D12’s description in example 1:

Example 2:

M12: [miso lepto . ]

[just a minute . ]

---

“and” (“the bucket”, “the side that you are standing at”). Even though “but the regular side” cannot be clefted in the larger sentence, it was still semantically and pragmatically linked to its preceding phrase (“but the regular side” is a clarification modifying “the side that you are standing at”), and thus the entire phrase was considered to be a single clause for our purposes.

<sup>2</sup> In Pair 12, the Director was seated at 0° and the Matcher was seated at 135°. By “regular side”, D12 meant the “side” of the circular table parallel to the room’s longer wall and perpendicular to the 180° heading (the canonical—“kanoniki”—side), rather than the “side” perpendicular to M12’s oblique 135° heading (“the side that you are standing at”).

## Appendix B

Comprehensive list of unique global strategies contributed by the Director and the Matcher in each pair.

Pair	Availability of Matcher's viewpoint	Structure Alignment	Global system introduced by Director	Global system introduced by Matcher
1	not Available	Neither	horizontal and vertical lines	zigzag system
2	Available	Matcher	none	none
3	not Available	Director	grid X-O system	none
4	Available	Director	none	none
5	not Available	Matcher	none	forming a reverse house, forming a rectangle
6	Available	Neither	axes (aligned with structure), quadrants, coordinates	axes (aligned with M6 & D6)
7	not Available	Matcher	forming a rectangle, lines	none
8	Available	Neither	quadrants, O-S system, forming a circle	vertical and horizontal segments
9	not Available	Director	symmetrical shape, lines, forming a reverse house	none
10	Available	Matcher	quadrants, semicircles	none
11	not Available	Neither	none	hexagon
12	Available	Director	cross	none
13	not Available	Director	cross, forming a rectangle	none
14	Available	Neither	forming a rectangle	none
15	not Available	Matcher	circumference	forming an X
16	Available	Director	forming a pyramid, house	circle
17	Available	Matcher	none	none
18	not Available	Matcher	forming a circle	locations according to a clock
19	not Available	Neither	forming a square	none
20	not Available	Director	horizontal, vertical lines, cross	none
21	Available	Director	quadrants, rectangle	none
22	Available	Neither	square	none
23	not Available	Neither	forming a square, locations according to a clock, forming a house	none
24	Available	Matcher	forming a rectangle	none



**References:**

- Anderson, A. H., M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert. (1991). The HCRC Map Task Corpus. *Language and Speech* 34, 351-366.
- Bavelas, J., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79, 941-952.
- Brennan, S. E. (1991). Conversation with and through computers. *User Modeling and User-Adapted Interaction*, 1, 67- 86.
- Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95-129). Cambridge, MA: MIT Press.
- Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. Ross (Ed.) *Psychology of Learning and Motivation*, 53, (pp. 301-345). Academic Press/Elsevier.
- Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics* 23, 13-32.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. Resnick, J. Levine, & S. Teasley, (Eds). *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: APA Books.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

- Degand, L., & Simon, A. C. (2009). On identifying basic discourse units in speech: theoretical and empirical issues. *Discours 4: Linearization and segmentation in discourse* (Special Issue), online: < <http://discours.revues.org/index5852.html> >.
- Duran, N. D. & Dale, R. (2014). Perspective-taking in dialogue as self-organization under social constraints. *New Ideas in Psychology*, 32, 131-146.
- Förster, J., & Higgins, E. T. (2005). How global versus local perception fits regulatory focus. *Psychological Science*, 16, 631-636.
- Fusaroli, R., Raczaszec-Leonardi J. & Tylén, K. (2013) Dialogue as interpersonal synergy, *New Ideas in Psychology*, 32, 147-157.
- Galati, A., & Avraamides, M. N. (2013). Flexible spatial perspective-taking: Conversational partners weigh multiple cues in collaborative tasks. *Frontiers in Human Neuroscience*. 7, 618. DOI:10.3389/fnhum.2013.00618
- Galati, A., & Avraamides, M. N. (2015). Social and representational cues jointly influence spatial perspective-taking. *Cognitive Science*. 39, 739-765.
- Garofalo, J., & Lester Jr, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, 163-176.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition*, 27, 181-218.
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathos (Ed.). *Everyday Language. Studies in Ethnomethodology* (pp. 97- 121). New York: Irvington.
- Graham, S.M., Joshi, A., & Pizlo, Z. (2000). The Traveling Salesman Problem: a hierarchical model. *Memory & Cognition* 28, 1191-1204.

- Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57, 596-615.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77-89.
- Heringer, H. J. (1993). Dependency Syntax - Basic Ideas and the Classical Model. In J. Jacobs, A. von Stechow, W. Sternefeld, and Th. Venneman (Eds). *Syntax – An International Handbook of Contemporary Research, Volume 1*. (pp. 298-316) Berlin - New York: Walter de Gruyter.
- Hölscher, C., Tenbrink, T., & Wiener, J. M. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition*, 121, 228-247.
- Horton, W. S., & Gerrig, R. J. (2002). Speaker's experiences and audience design: knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, 47, 589-606.
- Isaacs, E.A. & Clark, H.H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, 116, 26-37.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Kraut, R. E., Lewis, S. H., & Swezey, L. W. (1982). Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology*, 43, 718-731.
- Kuhlen, A. K., Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, 47, 567 – 587.

- Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Mainwaring, S. D., Tversky, B., Ohgishi, M., Schiano, D. J. (2003). Descriptions of simple spatial scenes in English and Japanese. *Spatial Cognition and Computation*, 3, 3-42.
- McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human spatial memory and navigation. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory*. Volume 2 in J. Byrne (Editor-in-chief), *Learning and memory: A comprehensive reference* (pp. 157-178). Oxford: Elsevier.
- Meilinger, T., Riecke, B.E., Bühlhoff, H.H. (2014). Local and global reference frames for environmental spaces. *Quarterly Journal of Experimental Psychology*, 67, 542-569.
- Mills, G. J. and Gregoromichelaki, E. (2010) *Establishing coherence in dialogue: sequentiality, intentions and negotiation*. In Proceedings of SemDial (PozDial). Adam Mickiewicz University, Poznan.
- Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: Listener effects on speakers' long-term memory for events. *Discourse Processes*, 26, 1-25.
- Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, 33, 1468- 1482.
- Schober, M.F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1-24.
- Schober, M.F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about location? *Discourse Processes*, 20, 219-247.
- Schober, M.F. (2009). Spatial dialogue between partners with mismatched abilities. In K.R. Coventry, T. Tenbrink, & J.A. Bateman (Eds.), *Spatial language and dialogue* (pp. 23-39). Oxford: Oxford University Press.

- Shockley, K., Richardson, D. C. and Dale, R. (2009), Conversation and Coordinative Structures. *Topics in Cognitive Science*, 1, 305–319.
- Taylor, H.A., Naylor, S.J., Faust, R.R., & Holcomb, P.J. (1999). Could you hand me those keys on the right? Disentangling spatial reference frames using different methodologies. *Spatial Cognition and Computation*, 1 (4), 381-397.
- Taylor, H.A. & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, 35, 371-391.
- Tenbrink, T. (2007). *Space, time, and the use of language: An investigation of relationships*. Berlin: Mouton de Gruyter.
- Tenbrink, T. (2015). Cognitive Discourse Analysis: Accessing cognitive representations and processes through language data. *Language and Cognition*. 7 (1), 98–137. DOI: 10.1017/langcog.2014.19
- Tenbrink, T., Andonova, E., & Coventry, K. (2008). Negotiating spatial relationships in dialogue: The role of the addressee. In J. Ginzburg, P. Healey & Y. Sato (Eds.), *Proceedings of LONDIAL*, pp. 201-208. London, UK: Queen Mary University.
- Tenbrink, T., Coventry, K. R., Andonova, E. (2011). Spatial strategies in the description of complex configurations. *Discourse Processes*, 48, 237-266.
- Tenbrink, T., Eberhard, K., Shi, H., Kübler, S., & Scheutz, M. (2013). Annotation of negotiation processes in joint-action dialogues. *Dialogue and Discourse* 4:2, 185-214.
- Tenbrink, T., & Wiener, J. (2009). The verbalization of multiple strategies in a variant of the traveling salesperson problem. *Cognitive Processing*, 10, 143-161.
- Tversky, B., Lee, P., & Mainwaring, S. D. (1999). Why do speakers mix perspectives? *Spatial Cognition and Computation*, 1, 399–412.
- Wiener, J. M., & Mallot, H. A. (2003). Fine-to-Coarse route planning and navigation in regionalized environments. *Spatial Cognition and Computation*, 3, 331 – 358.

## Footnotes

<sup>1</sup> In Galati & Avraamides (2015), the memory representations of Directors were assessed before the description phase in order to examine whether a priori information about the partner's viewpoint influenced how they organized spatial information in memory. The first memory test involved responding to a series of trials requiring Directors to make judgments of relative direction (JRDs) by indicating through the use of a joystick the location of objects from imagined perspectives ("Imagine you are at the bucket facing the candle. Point to the marble."); the second memory test involved reconstructing a drawing of the array. These memory tests were intended to examine the preferred direction or "organizing direction" (McNamara, 2003) by which Directors encoded the spatial array in memory. A consequence of using an organizing direction is that spatial relations from this direction can be retrieved from memory more readily (reflected in the orientation of the array drawings, and in facilitation in terms of accuracy and latency in the JRD task) compared to relations that have to be inferred.

According to the array drawings, virtually all Directors identified the relative locations of array objects correctly. Still, the organizing direction of memory—as reflected in both tasks--did differ across conditions, depending on the convergence of social and environmental cues. Directors who had studied the array while aligned with its intrinsic structure were more likely to use that axis as an organizing direction. Directors misaligned with the structure used their egocentric viewpoint more frequently as an organizing direction when not knowing their Matcher's viewpoint at study, but used the structure's axis more frequently when they knew the Matcher would be aligned with it.

## Figure captions

*Figure 1.* The seven-object array used, indicating 0°, 135°, and 225° viewpoints.

*Figure 2.* The proportion of BDUs containing references to a Global or a Local system, across the three conditions of structure alignment (Director-aligned, Matcher-aligned, Neither-aligned) and the two conditions of availability of the Matcher's viewpoint at study (Available, Unavailable).

*Figure 3.* The proportion of BDUs of each segment of the dialogue (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>) containing references to a system (a Global system or its components, or a Local system), across the two conditions of availability of the Matcher's viewpoint at study (Available, Unavailable).

*Figure 4.* The proportion of BDUs of each segment of the dialogue (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>) containing contributions of the Matcher that were classified as (i) Acceptances, (ii) Queries (i.e., Clarification and Expansion Requests), (iii) New Proposals, (iv) Recaps, (v) Responses to the Director's questions, and (vi) Metacomments.

Figure 1





Figure 2

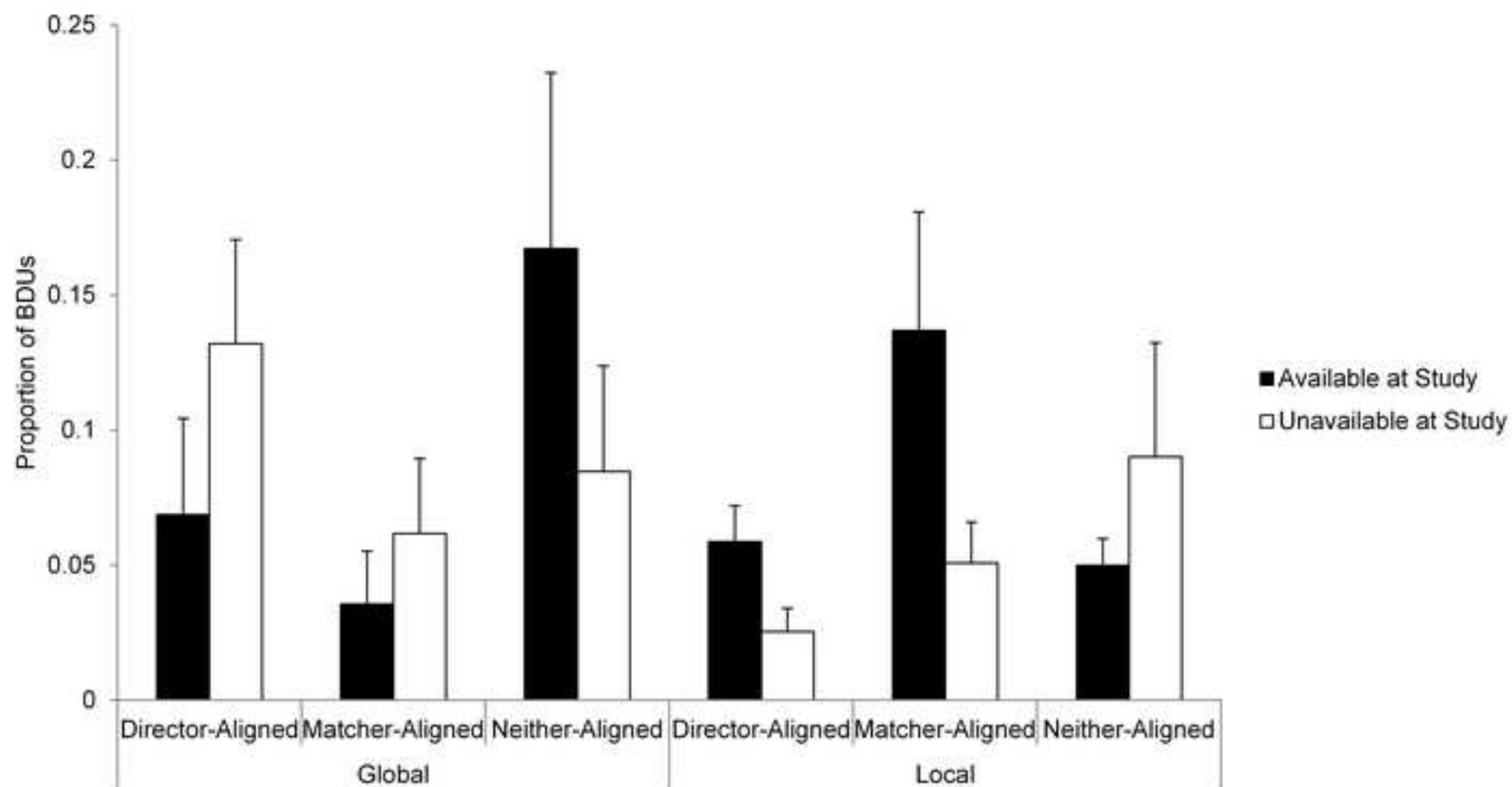


Figure 3

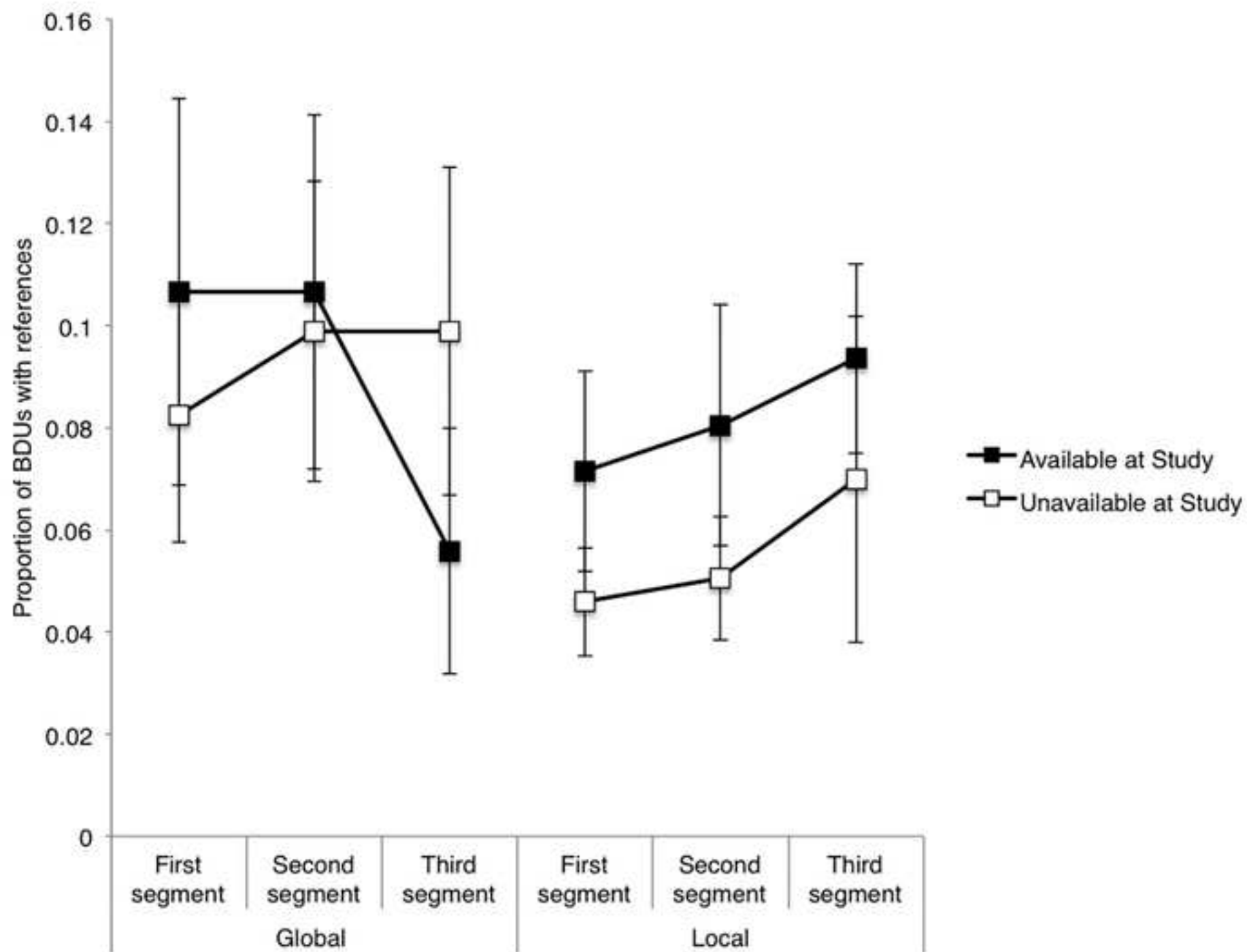


Figure 4

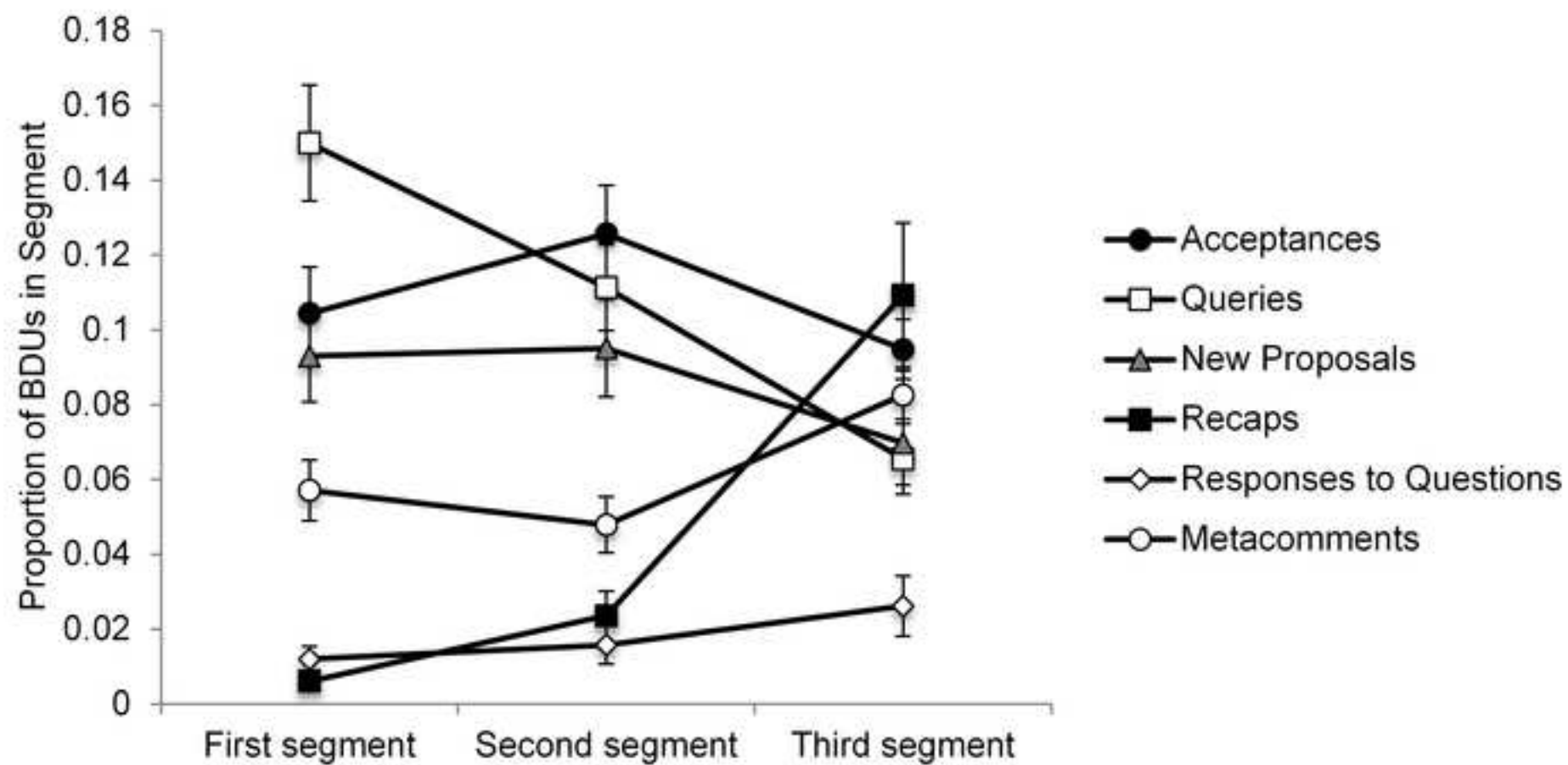


Table 1

*Table 1.* Excerpt of a dialogue, indicating the presence of prosodic and syntactic boundaries (marked as 1 or 0), whose convergence resulted in a Basic Discourse Unit (BDUs numbered under the “BDU” column). Speech associated with BDUs is contained in square brackets; it is transliterated from Cypriot Greek and translated in English below. Asterisks mark self-interruptions and angle brackets contain elongated phonemes. Gestures constituting a turn, or else critical to following the pair’s coordination process, are contained in brackets in italics. In Pair 1, the Director is seated at 225° and the Matcher at 135°; the Matcher’s viewpoint was not available at study.

Speaker	Prosodic	Syntactic	BDU	Speech
D1	0	1		[iha tris grammes orizonties [I had three horizontal lines
	1 (Self-interruption)	1	24	dj’ itan telia* ] and they were completely*]
M1	0	0		<i>(gestures 3 parallel lines with a sideways-oriented palm parallel to her torso, each gesture moving away from her body, visible over the barrier)</i>
D1	1 (Pause & falling intonation)	1	25	[bra<a>vo !] [exa<a>ctly!]
	0	1		[to “katheta sou” en’ to katheta sou<ou> [the “perpendicular to you” is perpendicular to you<ou>
	1 (Pause)	1	26	to katheta] the perpendicular]
M1	1 (Pause)	1	27	[n<n>e ] [yea<a>h] <i>(gestures with a sideways-oriented palm perpendicular to her torso, below the barrier)</i>
D1	0	1		[to lipon [so
	0			ara etsi opos ime ego tora therefore the way I am now
	1 (Pause)	1	28	an mporis na to skeftis ] if you can imagine that]

Table 2

Table 2. Excerpt from a dialogue illustrating the coding of global references (Global), references to global constituents (GI-constit), local references (Local), and the Matcher’s contribution (M-contribution). Dialogue is linearized in terms BDUs, with their associated speech contained in square brackets and numbered in the BDU column (transliterated from Cypriot Greek and translated in English below). The presence and absence of a global or local reference is indicated by 1 and 0, respectively, with a brief characterization of that system underneath in brackets. Each BDU in the Matcher’s turns is classified as a particular type. This example showcases an instance where the Matcher proposed a global system as a novel conceptualization (i.e., Expansion: new proposal) for apprehending the array.

Speaker	BDU	Global	GI-constit.	Local	M-contribution	Speech
M1	253	1 (zigzag)	0	0	Expansion: new proposal	[me alla logia eschematizetun kati san zigzag ? / ] [in other words it formed something like a zigzag ? /]
D1	254	0	0	0	-	[bra<a>vo !] [exa<a>ctly!]
M1	255	0	0	0	Acceptance	[mmm ] [mmm ]
D1	256	0	1 (zigzag lines)	0	-	[apla {to ena ant*} itan thkio antikimena mprosta thkio antikimena piso thkio antikimena mprosta dje ena {...}] [it’s just that {one obj*} there were two objects at the front two objects at the back two objects at the front and one {...}]
M1	257	0	0	1 (center referent)	Expansion: new proposal	[dje to<o> / kuvas mes ti mesi peripu] [and the<e> bucket in the middle approximately]
D1	258	0	0	0	-	[bra<a>vo] [exa<a>ctly]

*Table 3.* Summary of findings for each of the four main research goals.

\* indicates results where  $p < .05$ ; <sup>+</sup> indicates trends that were non-significant ( $.05 < p < .10$ )

### Research Question

1) How do spatial strategies (global and local) relate to contextual factors in the task?

2) How do spatial strategies (global and local) relate to the Matcher's contributions?

3) How does the distribution of spatial strategies and the Matcher's contributions change over time?

4) How do spatial strategies relate to task success?

### Main findings

1a) Pairs use more global than local strategies when:

- they know in advance that neither of them will be aligned with the intrinsic structure\*
- Directors aligned with the intrinsic structure don't know in advance their Matcher's viewpoint<sup>+</sup>

1b) Pairs use more local strategies when:

- Directors know in advance (vs. not) that the Matcher is aligned with the intrinsic structure\*

2a) Compared to other conditions, when Directors are aligned with the intrinsic structure, Matchers:

- use global and local strategies less frequently<sup>+</sup> (also supported by qualitative evidence) and make fewer queries<sup>+</sup>

2b) Matchers use:

- more metacomments as pairs make more global references<sup>+</sup>
- more recaps and clarification requests as pairs make more local references\*

3a) Over time, pairs:

- increase local references\*
- decrease global references in the final segment of the dialogue, but only when the Matcher's viewpoint had been known in advance<sup>+</sup>
- The complementary relationship between global and local references attenuates (holding only for in the first segment of the dialogue\*)

3b) Over time, Matchers:

- decrease their queries\* and new proposals\*
- increase their recaps\* and metacomments\* in the final segment of the dialogue

4a) As pairs increase their use of:

- global strategies, they reconstruct less distorted<sup>+</sup> and less rotated<sup>+</sup> arrays, especially earlier in dialogue (in the first<sup>+</sup> and second\* segment)
- local strategies, they reconstruct more distorted<sup>+</sup> arrays, especially earlier in dialogue (for the first segment\*)

4b) The Matcher's feedback predicts:

- reduced efficiency (increased turns\* and BDUs\*) as metacomments indicating confusion increase
- reduced accuracy as queries increase (lower  $BDr^*$ ) and as new proposals increase (higher rotation<sup>+</sup>)
- increased accuracy (less rotation\*, higher  $BDr^+$ ) as recaps increase